VIROLOGY 206, 403–412 (1995)

Coding Density of the Turnip Yellow Mosaic Virus Genome: Roles of the Overlapping Coat Protein and p206-Readthrough Coding Regions

KATHRYN L. BRANSOM, JOHN J. WEILAND,[1] CHING-HSIU TSAI,[2] AND THEO W. DREHER*,[3]

*Department of Agricultural Chemistry, Program in Genetics and *Center for Gene Research and Biotechnology, Oregon State University, Corvallis, Oregon, 97331*

Received August 10, 1994; accepted October 24, 1994

More than one-third of the turnip yellow mosaic virus (TYMV) genome simultaneously encodes two ORFs. We have investigated the functions of the overlapping coat protein ORF and readthrough domain of ORF-206 in the 3′ region of the genome. TYMC-206 RNA, in which a second stop codon has been positioned to prevent ORF-206 readthrough, induced infections in protoplasts and plants that were indistinguishable from wild type. ORF-206 readthrough is thus nonessential. Nevertheless, TYMV-221 RNA, in which the ORF-206 stop codon was replaced with a tyrosine codon to force readthrough, was infectious to protoplasts, suggesting that a role for ORF-206 readthrough under certain conditions is possible. TYMV RNA variants that produce truncated or no coat protein were used to show that the coat protein is dispensable for local movement but necessary for systemic spread of virus in plants. Studies in protoplasts showed that (−) RNA levels are normal in the absence of coat protein, but (+) strand levels are decreased about 10-fold relative to wild-type infections. A mutant with a short C-terminal coat protein extension that formed virions less stable than normal demonstrated the protective role of capsids toward genomic RNA. The evolutionary implications of the dense information content of the TYMV genome are discussed.

INTRODUCTION

Among the alphavirus-like supergroup of positive strand viruses (Goldbach *et al.,* 1991), turnip yellow mosaic virus (TYMV) and the other tymoviruses stand out in possessing genomes of unusually compact coding structure. Of the 6318 nucleotides in the TYMV genome, only the 5′ 87 nucleotides and the 3′ 108 nucleotides are noncoding. More remarkably, 2283 nucleotides (36.1% of the genome) of the 6123 nucleotide coding region (96.9% of the genome) simultaneously participate in the coding of two extensively overlapping open reading frames (ORFs; Morch *et al.,* 1988; Keese *et al.,* 1989; Dreher and Bransom, 1992). The most extensive overlap involves a stretch of 1877 nucleotides in the 5′ region of the genome, and the dual coding role of these nucleotides is well understood. These nucleotides are involved in encoding both p69 (ORF-69), a protein that is necessary for cell-to-cell movement of the virus (Bozarth *et al.,* 1992), and p206 (ORF-206), which contains helicase-like and polymerase-like domains that are essential for RNA replication (Weiland and Dreher, 1993; Fig. 1). Proteins p69 and p206 are initiated with approximately equal efficiency in rabbit reticulocyte lysates at AUG codons only 7 nucleotides apart (Weiland and Dreher, 1989). p206 is a polyprotein that undergoes autoproteolysis to yield N-terminal 150-kDa and C-terminal 70-kDa proteins (Bransom *et al.,* 1991; Fig. 1).

In this paper we address the dual coding role of the second extensive overlap in the TYMV genome, a stretch of 406 nucleotides downstream of ORF-206 (nucleotides 5644–6049). These nucleotides simultaneously encode most of a potential readthrough extension of ORF-206 (nucleotides 5627–6049) and the N-terminal two-thirds of the 20-kDa coat protein, which is encoded by ORF-CP (nucleotides 5644–6210; Fig. 1). Readthrough of the single UAG (amber) codon that terminates ORF-206 occurs *in vitro* in both rabbit reticulocyte and wheat germ extracts (Morch *et al.,* 1986; K.L.B. and T.W.D., unpublished observations), and results in the extension of p206 to produce a 221-kDa protein (p221). It is not known whether readthrough occurs *in vivo* or whether it plays a role in viral replication. The coat protein is expressed from a subgenomic RNA that corresponds to sequences between nucleotide 5625 and the end of the genomic RNA (Guilley and Briand, 1978). The subgenomic RNA thus partially overlaps the end of ORF-206, and it has been suggested that a conserved "tymobox" sequence at the end of ORF-206 may function as the subgenomic promoter (Ding *et al.,* 1990). It is thus likely that subgeno-

[1] Present address: Department of Plant Pathology, North Dakota State University, Fargo, ND 58105.

[2] Present address: Agricultural Biotechnology Laboratories, National Chung Hsing University, Taichung, 40227, Taiwan.

[3] To whom correspondence and reprint requests should be addressed. Fax: (503) 737-0497.

0042-6822/95 $6.00

mic transcription control signals overlap coding regions, a further example of the unusually dense information content of the TYMV genome.

We report here our studies on the significance of the downstream overlapping coding region of TYMV by genetically investigating the roles of the CP-ORF and the ORF-206 readthrough domain.

MATERIALS AND METHODS

Plant materials and virus stocks

Turnip (*Brassica rapa* cv. Just Right) and Chinese cabbage (*Brassica pekinensis* cv. Spring-A1) were grown in a controlled chamber at 21° with 16-hr daylength. Plasmid pTYMC, a cDNA clone from which infectious genomic RNA can be transcribed with T7 RNA polymerase, has been described previously (Weiland and Dreher, 1989). We refer to transcripts and virus derived from this wild-type cDNA clone as TYMC RNA and TYMC, respectively. Virus and viral RNA were prepared by standard methods (Lane, 1986).

Mutant plasmids, *in vitro* transcription, and *in vitro* translation

Construct pTYMCΔ5707–6062 was made by deleting sequences between the *Pvu*II[5708] and *Sma*I[6062] restriction sites of pTYMC (Weiland and Dreher, 1993; Fig. 1). Plasmid pTYMC-5985[fs] was constructed by end-filling *Bgl*II-digested pTYMC DNA with DNA polymerase I (Klenow fragment), followed by religation of the blunt ends; dideoxynucleotide sequencing revealed that 5 bp were inserted into this clone instead of the expected 4 bp. The construction of pTYMC-C107/G96 (Tsai and Dreher, 1992) and pTYMC-178/224 (Bozarth *et al.*, 1992) have been described previously. Plasmids pTYMC-noCP, pTYMC-221, and pTYMC-206 were constructed by PCR mutagenesis according to the "megaprimer" method of Sarkar and Sommer (1990). Mutagenic (+)-sense oligonucleotides were used in conjunction with a (−)-sense oligonucleotide spanning the 3′ *Hin*dIII site of pTYMC in a primary PCR reaction using linearized pTYMC as template. After phenol/chloroform extraction and removal of the nontemplate-specified 3′ adenosine residues by incubation with 2.5 units of the Klenow fragment of DNA polymerase I (15 min at room temperature), the DNA fragment containing the desired mutation was used as primer in a secondary PCR reaction, in conjunction with the (+)-sense oligonucleotide that primed at position 4838 on the TYMC genome. This allowed the mutation-bearing fragment to be cloned into pTYMC via the *Pst*I[5404] and *Hin*dIII[6318] sites. The entire subcloned sequence was verified by dideoxy sequencing.

DNA templates (pTYMC and the mutant derivatives) were linearized at the unique *Hin*dIII site and were transcribed in the presence of cap analogue as reported previously (Weiland and Dreher, 1989). Transcript RNAs were translated in a rabbit reticulocyte lysate as described by Bransom *et al.* (1991).

Protoplast and plant inoculations

Protoplasts were prepared from turnip plants, inoculated (4×10^5 cells/inoculation), and subsequently incubated under constant light at 25° for 48 hr, as reported previously (Weiland and Dreher, 1989). Viral (1 μg) or capped transcript RNA (5 μg) were used both to inoculate protoplasts and for mechanical inoculations of Chinese cabbage and turnip plants. For plant inoculations, 5 μl of inoculum was diluted with 5 μl of 2× inoculation buffer (100 m*M* glycine, 60 m*M* K_2HPO_4, pH 9.2, and 6 mg/ml bentonite) and rubbed onto the outer leaves of 4-week-old Chinese cabbage or turnip plants.

Tissue extraction, Western and Northern blotting

For analysis of individual lesions, a 10-mm leaf disc was punched with the lesion at the center of the disc. Petiole samples, 1 cm in length taken from the base of the leaf, were harvested from inoculated leaves. Because of variations in lesion and petiole development, these analyses are meant to identify large differences in accumulations of viral products and are not suited for quantitative analysis. To sample young expanding leaves above the inoculated leaf, leaf punches (1-cm diameter) were made 14 days after inoculation, when wild-type symptoms were well developed. Plant tissue was ground in liquid nitrogen and then ground again in 200 μl of tissue extraction buffer (70 m*M* Tris–HCl, pH 7.2, 2 *M* urea, 1% SDS, 0.6% β-mercaptoethanol, 2 m*M* EDTA, 1 m*M* PMSF, and 3 mg/ml bentonite). The samples were vortexed thoroughly, centrifuged for 4 min at 4°, and a portion (150 μl) was removed. Nucleic acids were collected from this 150 μl fraction after two phenol/chloroform extractions. The rest of the tissue sample (50 μl) was boiled for 5 min and then centrifuged for 15 min at 4° prior to analysis of proteins in the supernatant via Western blotting.

Protoplasts (2×10^5) were harvested for Northern blot analysis by vortexing in 200 μl of protoplast extraction buffer (200 m*M* Tris–HCl, pH 8.0, 100 m*M* EDTA, 125 m*M* NaCl, 1% SDS and 3 mg/ml bentonite), and immediately extracting with phenol/chloroform prior to precipitation of the nucleic acids with ethanol. An equal number of protoplasts were harvested for Western blotting. The levels of viral coat protein in tissue and protoplast extracts were analyzed as described previously (Weiland and Dreher, 1989), except that enzyme-linked chemiluminescence (Amersham) was used for detection.

Nucleic acids harvested from tissue samples or protoplasts were glyoxalated, electrophoresed in 1% agarose gels, and alkaline-blotted to nylon membranes as de-

scribed (Weiland and Dreher, 1989; Vrati *et al.,* 1987). Hybridization was done in the presence of 0.2 mg/ml polyanetholesulfonic acid (Calbiochem). Riboprobes complementary to either the 3′-terminal 259 bases of the TYMC genomic RNA or the 5′-terminal 259 bases of (−)-sense RNA were synthesized with T7 RNA polymerase, and hybridization was performed at 65° for 12–16 hr. The membrane was washed twice (7.5 m*M* Tris–HCl, pH 8.0, 37.5 m*M* NaCl, 0.5 m*M* EDTA, 0.1% SDS, 0.1% sodium pyrophosphate) for 15 min at 65° prior to exposure to preflashed X-ray film. Laser scanning densitometry of exposed films permitted quantitative analysis.

cDNA synthesis, PCR amplification, and sequencing of progeny RNAs

Viral RNA was isolated after serial passages through Chinese cabbage or turnip plants and used as template for the synthesis of cDNA, using a phosphorylated (−)-sense oligonucleotide primer. Products of the cDNA reaction were subjected to PCR amplification directly, without purification, using a phosphorylated (+)-sense oligonucleotide primer, and without additional (−)-sense primers. The PCR products were purified from a low-melting point agarose gel prior to treatment with the Klenow fragment of DNA polymerase I to remove nontemplate-derived 3′ adenosine residues. Blunt-ended PCR products were then cloned into *Sma*I-linearized pUC18, and plasmid clones containing inserts were sequenced according to the method of Chen and Seeburg (1985).

Electron microscopy

Leaf tissue was fixed by vacuum infiltration with 1% glutaraldehyde in 0.2 *M* phosphate buffer, pH 7.2. The fixed tissue was squashed to extrude fixed material including virions, stained with phosphotungstic acid, and examined on a Formvar-coated grid in a Philips CM12 scanning-transmission electron microscope at the OSU EM center.

RESULTS

Replication of mutants with altered ORF-206 readthrough

In order to examine the role of the TYMV ORF-206 readthrough domain, two mutants were created to either prevent expression of this domain (TYMC-206) or to force constitutive readthrough, so that p221 rather than p206 is produced from ORF-206 (TYMC-221). In TYMC-206, an additional in-frame termination codon (UGA) was placed two codons downstream of the UAG codon that terminates ORF-206. The single C5634 → G mutation that creates the UGA codon (Fig. 1) falls at position 10 in the 5′ untranslated region of the coat protein subgenomic RNA (the coat protein initiation codon is at position 20–22) and should not affect coat protein expression. TYMC-221 has a tyrosine codon (UAC) in place of the UAG termination codon (Fig. 1). A tyrosine codon was chosen on the basis of the suggestion that $tRNA^{Tyr}$ promotes UAG readthrough during translation of tobacco mosaic virus (TMV) RNA (Beier *et al.,* 1984). TYMC-221 RNA additionally has an A5626 → G mutation that assisted mutant detection during cloning; this mutation replaces the CAA glutamine codon with the CAG glutamine codon at the 3′ end of ORF-206. The two mutations in TYMC-221 RNA alter the 5′ end of the coat protein subgenomic RNA from AAUAGCAA. . . to AGUACCAA. . . , and might thus affect the efficiency of subgenomic RNA synthesis. Translation of TYMC-206 and TYMC-221 RNAs in a rabbit reticulocyte lysate produced ORF-206-encoded proteins of the predicted size (Fig. 2).

TYMC-206 and TYMC-221 RNAs were tested for their ability to replicate in turnip protoplasts. Western blot analysis of protoplast extracts harvested 48 hr after inoculation showed that TYMC-206 RNA supported coat protein accumulations similar to wild-type TYMC RNA (Fig. 3A, lanes 3 and 4), while TYMC-221 RNA supported lower levels, 8% those of wild type (Fig. 3A, lane 2). Analysis of RNA products via Northern blots showed that the accumulations of TYMC-206 genomic and subgenomic (+)-strands (Fig. 3B, lane 4) and genomic (−)-strands (Fig. 3C, lane 4) were similar to those of wild-type TYMC RNA. TYMC-221 (+)-sense RNAs accumulated to levels well below those of TYMC infections (Fig. 3B, lane 3); (−)-sense RNAs were not detected from TYMC-221-inoculated protoplasts (not shown). The presence of coat protein in TYMC-221-inoculated protoplasts demonstrates unequivocally that some replication of TYMC-221 RNA and transcription of subgenomic RNA has occurred.

The ability of TYMC-206 and TYMC-221 RNAs to establish an infection on Chinese cabbage and turnip plants was also tested. On both hosts, TYMC-206 RNA produced infections that were indistinguishable from those produced from wild-type TYMC RNA. Symptom development occurred with the same timing (lesions after 5 days, systemic symptoms after 7 days), and virus yields were approximately equivalent (≈1 mg/g tissue) over several passages. Western analysis with anti-TYMV antiserum showed a similar accumulation of coat protein in tissues of plants inoculated with TYMC-206 RNA, as compared to plants inoculated with TYMC RNA (Fig. 4, lanes 3 and 4). TYMC-206 RNA was present in high yield in lesions and systemically infected tissue (Figs. 5A and 5B, lane 2) and was recovered unchanged from turnip plants after four sequential passages (inoculum consisted of either sap or viral RNA from mature symptomatic leaves), as determined by PCR amplification, cloning, and DNA se-

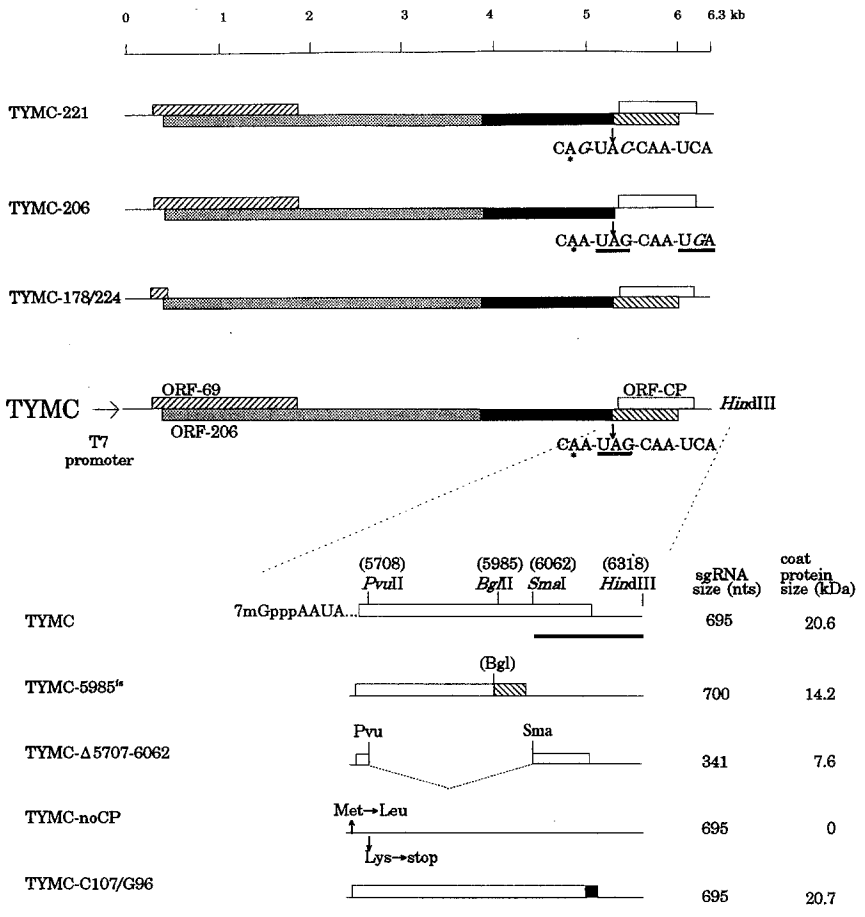

FIG. 1. Diagrams of the TYMV genome and mutants used in this study. Map of the TYMC genome (middle of figure) showing the ORF-69 and ORF-206 overlap in the 5′ region, and the ORF-CP (coat protein) and ORF-206 readthrough overlap in the 3′ region. ORF-206 consists of 3 domains; the first two comprise the coding region of the p206 polyprotein, which is proteolytically cleaved to yield N-terminal p150 (stippled) and C-terminal p70 (solid) proteins. Readthrough of the UAG stop codon (positioned at the arrow and underlined in the sequence shown below the diagram) permits translation of the ORF-206 readthrough domain (hatched). An asterisk marks the initiation site of the coat protein subgenomic RNA. The T7 promoter and 3′-*Hin*dIII site used to linearize pTYMC templates for transcription are shown. The changes present in mutant genomes are shown above and below the diagram of wild-type TYMC. Single nucleotide substitutions are marked in italics (TYMC-221 and TYMC-206), and restriction sites in pTYMC used in constructions are indicated. The lower part of the figure shows mutants used to study the role of the coat protein. The changes are shown in the context of the subgenomic RNA, with the open box representing ORF-CP; the overlapping ORF-206 readthrough domain is not marked. Predicted sizes of the sgRNAs and coat proteins resulting from the mutations are shown. The solid bar under the TYMC diagram spans sequences present in the riboprobes used in Northern blot studies (*Sma*I^{6062}–*Hin*dIII^{6318}).

quencing of the 3′ noncoding region (four separate clones were sequenced; data not shown). In contrast, inoculation with TYMC-221 RNA failed to produce symptoms on 10 different plants, even after extended periods of time (8 weeks). No local lesions appeared. Western and Northern analyses of plant tissue extracts failed to detect any viral products in these plants, even in inoculated leaves (Fig. 4, lane 2; Fig. 5, lane 1).

Replication in protoplasts of TYMC derivatives with mutant coat protein genes

Two TYMC derivatives that produce shortened forms of the 189 amino acid 20-kDa coat protein were studied for their effect on replication and symptom formation: a deletion mutant lacking a large central region of the coat protein gene, and a frameshift mutant encoding a truncated coat protein but lacking no RNA sequences (Fig. 1). The coat protein ORF of TYMC-Δ5707–6062 lacked sequences between the *Pvu*II^{5708} and *Sma*I^{6062} restriction sites of pTYMC. This in-frame deletion results in the loss of codons 22–139, leaving a 71-codon ORF encoding a coat protein of expected MW 7.6 kDa. The RNA frameshift mutant TYMC-5985^{fs} had 5 additional nucleotides inserted at the *Bgl*II^{5985} site to produce a frameshift, resulting in a coat protein ORF with wild-type codons 1–114 fused to 21 codons encoded by the readthrough domain of ORF-206 (refer to Fig. 1). This mutant ORF terminates at codon UGA^{6050} and encodes a hybrid coat protein of expected MW 14.2 kDa (Fig. 1). A third coat protein mutant (TYMC-noCP), which harbors a mutation in the initiation codon and prevents translation of the

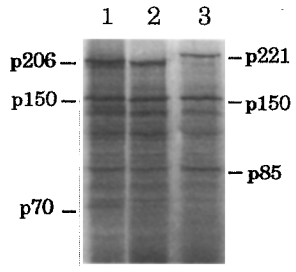

FIG. 2. Cell-free translations of TYMC and mutant RNAs. RNAs transcribed *in vitro* were translated in a rabbit reticulocyte lysate in the presence of [^{35}S]methionine and separated by SDS–PAGE (7%). The migration positions of p206 and the readthrough product p221 are indicated. Note that the proteolytic cleavage products of p206 (p150 and p70, indicated at left) correspond to cleavage products of 150 and 85 kDa, respectively, when translational readthrough is forced (indicated at right of lane 3). Lane 1 shows a translation of TYMC RNA, lane 2 of TYMC-206 RNA, lane 3 of TYMC-221 RNA.

coat protein ORF was also studied; the initiator methionine was changed to a Leu codon (UUG), and a termination codon (UAA) was positioned 5 codons further downstream (altered from an AAA Lys codon). The initiation codon mutation did not change the coding sequence of the ORF-206 readthrough domain, but the termination codon mutation resulted in an Asn to Lys substitution at amino acid 1855 in the putative 221-kDa protein. A mutant with a defective movement protein that replicates normally in protoplasts (TYMC-178/224, Fig. 1; described by Bozarth *et al.*, 1992) was used as a control for monitoring cell-to-cell movement in whole plants.

Northern blot analysis of extracts made from turnip protoplasts 48 hr after inoculation showed that all three coat protein mutant genomes were able to replicate in protoplasts (Figs. 3B and 3C). Accumulations of (+)- and (−)-strand genomic and of subgenomic RNAs were detected in each case. However, the levels of (+)-sense RNAs were decreased relative to wild-type TYMC by about a factor of 10 (Fig. 3B, lanes 5 and 7–9). In contrast, the accumulations of (−)-strand genomic RNA were similar between wild-type TYMC, TYMC-noCP, TYMC-Δ5707–6062, and TYMC-5985^{fs} infections (Fig. 3C, lanes 5–8). Detection of the mutant coat proteins with anti-TYMV antiserum in Western blots was not useful for quantitation due to the missing epitopes normally found in the wild-type coat protein. Nevertheless, TYMC-Δ5707–6062 did produce a relatively strong signal corresponding to a protein of the expected MW of 7.6 kDa (Fig. 3A, lane 8). These results indicate that the TYMV coat protein does not have an essential role in viral replication in protoplasts, but does appear to influence the accumulation of (+)-sense genomic and subgenomic RNAs.

The replication of a fourth mutant with an altered coat protein gene was also studied. Mutant TYMC-C107/G96 (Fig. 1) encodes a coat protein with a C-terminal extension of five amino acids (encoding Tyr–Val–Leu–Asp–Arg) due to a substitution in the normal termination codon of the coat protein ORF ($^{-109}$UAA^{-107} replaced by $^{-109}$UAC^{-107}; negative numbering to indicate distance

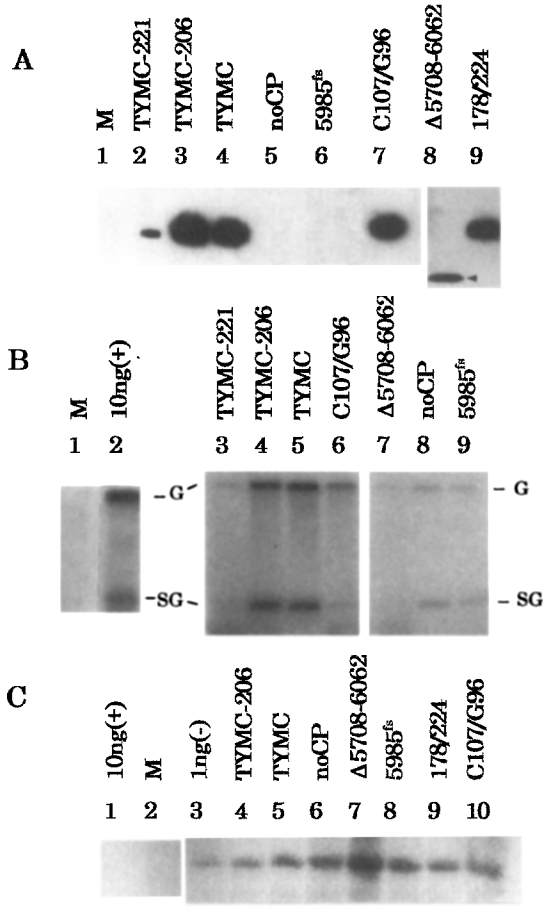

FIG. 3. Accumulation of viral products in turnip protoplasts inoculated with mutant forms of TYMC RNA. The indicated transcript RNAs (5 μg) were inoculated onto 4 × 10^5 turnip protoplasts and incubated for 48 h at 25° before harvest. (A) Western blot analysis to detect viral coat protein. Protein extracts corresponding to 5 × 10^4 protoplasts (except in lane TYMC-221, which contains 1 × 10^5 cells) were electrophoresed on 12% SDS–PAGE gels prior to transfer to nitrocellulose and chemiluminescent immunodetection with anti-TYMV antiserum. The position of the truncated coat protein encoded by TYMC-Δ5707–6062 is indicated with an arrowhead. Lane M refers to a mock inoculation. (B) Northern blot to detect (+)-sense genomic and subgenomic RNAs. Total nucleic acids from 1 × 10^5 cells were Northern blotted, hybridized to a (−)-sense ^{32}P-labeled RNA probe (complementary to nucleotides 6063 to 6318), and the blot was autoradiographed. M refers to a mock inoculation. A control lane had 10-ng virion (+) RNAs loaded (lane 2). G and SG indicate the migration positions of genomic and subgenomic RNAs, respectively. The blot was exposed for 2 days without an intensifying screen. (C) Northern blot analysis to detect (−)-sense RNAs. Cell extracts were analyzed as in (B) except for the use of a (+)-sense RNA probe (TYMC 3′-end sequences from nucleotides 6063 to 6318). The blot was exposed for 4 days with an intensifying screen. The signal for Δ5707–6062 is unusually strong in this experiment. Control lanes had 10-ng (+) virion RNA (lane 1) or 1-ng genomic (−) strand transcripts loaded (lane 3).

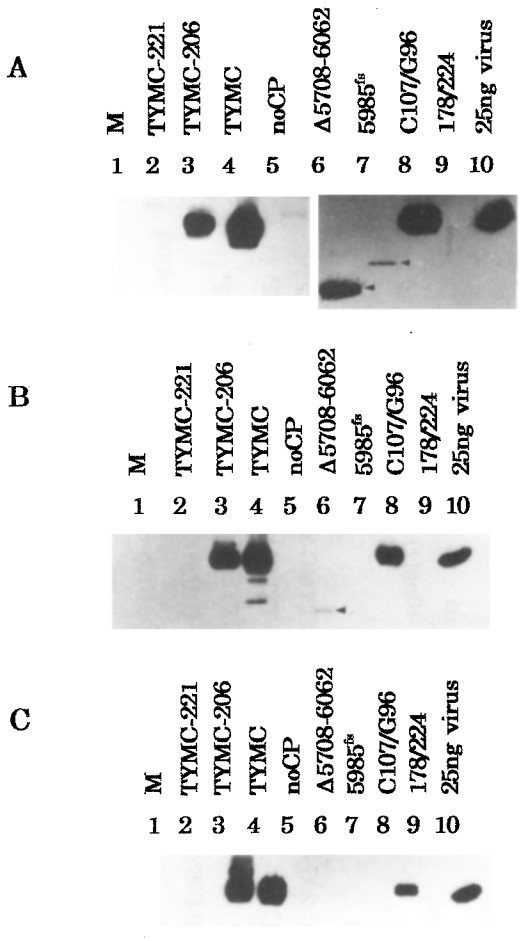

FIG. 4. Western blot analysis of coat protein accumulations in turnip plants inoculated with mutant forms of TYMC RNA. Transcript RNAs (5 μg) were inoculated onto 4-week-old turnip plants and tissues were harvested for extraction 14 days postinoculation. Denatured extract corresponding to 0.1-mg fresh weight was analyzed as described in the legend to Fig. 3A. The deleted forms of coat protein encoded by mutant RNAs are indicated by arrowheads. (A) Extracts from individual lesions on the inoculated leaf (or inoculated region in case of TYMC-178/224). (B) Tissue extracted from the petiole of the inoculated leaf. Note that Δ5707–6062 CP was detected in all inoculated plants, while 5985^{fs} CP was detected in two of seven plants. (C) Tissue extracted from a young expanding leaf above the inoculated leaf. Lanes labeled M refer to a mock inoculation. Lanes 10 were loaded with 25 ng of TYMV virions.

from the 3′-CCA end of TYMC RNA). This mutant originated from a study of second site mutations that suppress deleterious mutations in the 3′ noncoding region of TYMC RNA (Tsai and Dreher, 1992). In the work reported here, mutant TYMC-C107/G96 replication was studied in both Chinese cabbage (Tsai and Dreher, 1992) and turnip protoplasts harvested 48 hr after inoculation. As determined from Northern blots, (+)-strand genomic and subgenomic RNA accumulations relative to wild type were 0.2–0.4 (Fig. 3B), while the relative levels of (−)-strand genomic RNA were similar to wild type (Fig. 3C, lanes 5 and 10). Western blots indicated that the mutant coat protein accumulated to levels of 0.5–1.0 relative to wild type (Tsai and Dreher, 1992; Fig. 3A, lanes 4 and 7).

Role of coat protein in the spread of infection in whole plants

Inoculation of Chinese cabbage or turnip plants with TYMC-noCP, TYMC-Δ5707–6062, or TYMC-5985^{fs} RNAs resulted in the appearance of local lesions on the inoculated leaves within the same time frame as inoculations with TYMC RNA (5–7 days), but no systemic symptoms appeared. Lesions induced by the mutant RNAs were smaller and sometimes showed a necrotic boundary. Lesions were individually harvested and extracted in order to examine the accumulation of viral products. Western blots of these extracts showed the presence of shortened coat proteins of a size consistent with the alterations made in the coat protein (except for mutant TYMC-noCP, which produces no coat protein) (Fig. 4A, lanes 5–7). Viral (+)-strand genomic and subgenomic RNAs were detected in Northern blots (Fig. 5A, lanes 4–6); the levels of all the mutant RNAs were well below wild type, espe-

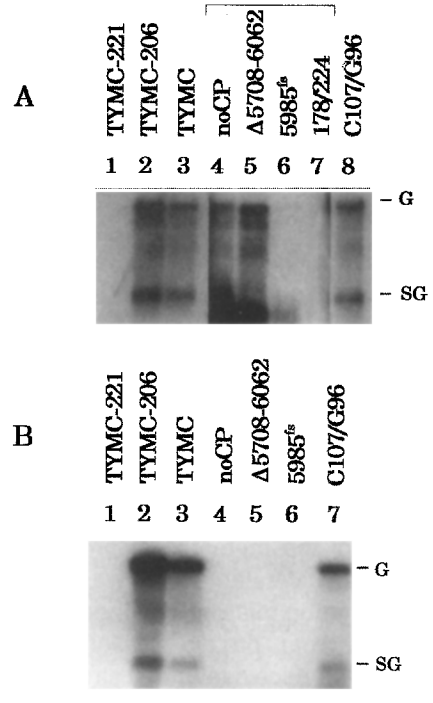

FIG. 5. Northern blot analysis of (+)-sense viral RNAs in turnip plants inoculated with mutant forms of TYMC RNA. RNA extracts were made from the same tissue samples analyzed in Fig. 4. Total nucleic acids were extracted, and a portion corresponding to 5 mg fresh tissue was analyzed on each lane as in Fig. 3B. G and SG indicate the migration positions of genomic and subgenomic RNAs, respectively. (A) RNAs extracted from individual lesions on the inoculated leaf. Lanes 4–7 (bracketed) were exposed 3 days with an intensifying screen, while other lanes were exposed for 7 hr with no screen. (B) RNAs extracted from young expanding leaf tissue taken above the inoculated leaf.

cially TYMC-5985^{fs}. Detection of viral products in local lesions indicates that cell-to-cell spread was occurring, since no products were detectable (nor lesions visible) in leaves inoculated with TYMC-178/224 RNA, which has a defect in the expression of the p69 movement protein (Bozarth *et al.,* 1992; Fig. 4A, lane 9; Fig. 5A, lane 7).

No systemic symptoms were seen on 10 or more plants displaying local lesions from TYMC-noCP, TYMC-Δ5707–6062, or TYMC-5985^{fs} inocula, nor were viral RNAs detected in tissue distant from the inoculation sites (Fig. 5B). Coat protein could also not be detected in young, expanding leaves (Fig. 4C) of turnip plants infected with these mutants, but low levels of coat protein were present in the petiole of leaves inoculated with TYMC-Δ5707–6062 RNA, and in the petiole of some plants inoculated with TYMC-5985^{fs} RNA (Fig. 4B, lanes 6 and 7). These results demonstrate that cell-to-cell movement can occur in the absence of coat protein, but that intact coat protein is required for efficient long-distance movement of TYMV.

Inoculation of plants with mutant TYMC-C107/G96

Chinese cabbage and turnip plants inoculated with TYMC-C107/G96 RNA developed normal systemic symptoms with the same timing as TYMC-inoculated plants. This mutant was recovered unchanged from Chinese cabbage plants after two sequential passages, as determined by PCR amplification of the 3′ noncoding region and DNA sequencing (Tsai and Dreher, 1992). However, repeated attempts to purify virions from these leaves by a procedure involving extraction at pH 4.8 and precipitation of virions with polyethylene glycol (Lane, 1986) were unsuccessful. The levels of viral products in symptomatic leaves harvested 2 weeks after inoculation were studied with Western and Northern blots, which showed the presence of significant levels of coat protein throughout the plant (Fig. 4, lanes 4 and 8) and of viral RNA in lesions and young expanding leaves (Fig. 5). In young systemically infected leaves, coat protein levels averaged 0.32 relative to wild type, while (+)-strand genomic RNA levels averaged 0.25 relative to wild type.

Because the above results suggested that TYMC-C107/G96 may form altered or unstable virions as a result of an interference in capsid assembly by the C-terminal extension, squash preparations were made from mature symptomatic leaves infected with either the mutant or wild-type viruses and fixed for electron microscopic inspection. Spherical virions were observed in each case, although fewer were present in the squash prepared from tissue infected with TYMC-C107/G96 (not shown). No features distinguishing wild-type and mutant virions were evident. The highly stable 28-nm icosahedral TYMV particles form from 180 subunits of the coat protein in the presence or absence of encapsidated RNA (Matthews, 1991), and it is not known whether the observed mutant capsids contained viral RNA. Stability differences between wild-type and mutant virions were also assessed by analyzing the levels of viral products as systemically infected leaves aged. While the relative coat protein levels remained fairly constant, there was a marked decline in TYMC-C107/G96 (+)-sense genomic RNA: RNA levels were 0.25 relative to wild type in leaves 2 weeks after inoculation, declining to 0.05 in the same leaves 2 weeks later (not shown).

Our observations indicate that apparently normal virions are present in TYMC-C107/G96 infections, but that they are less stable than normal. They cannot be recovered by normal procedures and have a diminished ability to protect virion RNA from degradation in infected leaves. The decrease of viral RNA but not of coat protein levels as mature infected leaves aged clearly demonstrates the role of virions in protecting the viral genome.

DISCUSSION

ORF-206 has a readthrough domain with no known role

Analysis for a potential role in TYMV replication of C-terminally extended ORF-206 proteins has involved the use of mutants which express alternate forms of the ORF-206 readthrough domain. The normal phenotype of TYMC-206 (Figs. 3, 4, and 5) demonstrates that the readthrough domain does not play a role in viral replication in protoplasts or plants, nor in symptom determination in plants. Interestingly, replacement of the ORF-206 termination codon with a Tyr codon to force p221 synthesis results in a viable (albeit attenuated) virus (TYMC-221). TYMC-221 replicated to reduced levels in protoplasts and was unable to systemically infect plants. Nevertheless, an active form of the viral RNA polymerase (the p70 analogue and/or p221; see Fig. 1) must have been present, indicating that the C-terminal extension does not radically alter the folding of the p70 domain. Thus, while ORF-206 readthrough is not an essential feature of TYMV gene expression, it may conceivably serve a role under special conditions not tested in our experiments, such as vector transmission, host range determination or replication in plants during unusual (e.g., stress) conditions.

It is also possible that readthrough occurs at a low level *in vivo,* even though it is not required. The sequence context of the ORF-206 termination codon (CAA-**UAG**-CAA-UCA) differs by only one base from the leaky CAA-**UAG**-CAA-U*U*A sequence from TMV that has been shown to support 2–5% readthrough in tobacco cells (Skuzeski *et al.,* 1990). The two codons following the termination signal were found to be important in promoting readthrough, and the TYMV context permitted readthrough at 40% the level of the TMV sequence (Skuzeski

et al., 1991). The presence in the TYMV genome of a potentially leaky termination codon leading into an extensive readthrough domain suggests that these features conferred a selective advantage during at least one stage of viral evolution. Since neither the termination codon identity or context, nor the presence of an extensive readthrough domain overlapping the coat protein gene, are conserved among the other tymoviruses (Srifah *et al.,* 1992), readthrough to produce alternative replication proteins does not appear to be a general feature of tymoviral gene expression.

CP is dispensable for cell-to-cell movement but necessary for systemic spread

The dispensability of ORF-206 readthrough has permitted a clear analysis of the role of the coat protein in TYMV infections. Using mutant TYMV genomes that express shortened forms of the coat protein or no coat protein at all, we have demonstrated that the expression of coat protein is not required for RNA replication in protoplasts or plants. However, the spectrum of progeny RNAs was somewhat altered, with lower accumulations of (+)-strand RNAs (both genomic and subgenomic), although there were normal accumulations of (−)-strand genomic RNA (Figs. 3B and 3C). Similarly altered ratios of (+):(−)-strand RNAs have been reported for infections caused by a number of mutant alphavirus-like plant viruses expressing either no coat protein or drastically altered coat proteins (Ishikawa *et al.,* 1991; Marsh *et al.,* 1991; Suzuki *et al.,* 1991; van der Kuyl *et al.,* 1991; Chapman *et al.,* 1992). It is probable that degradative loss of (+)-strand RNAs, normally protected by encapsidation, partly accounts for these observations. It is also possible that the coat protein is directly involved in regulating (+)-strand accumulation at the level of RNA synthesis, but additional experimentation will be necessary to determine if this occurs during TYMV infection.

Plant inoculation experiments with mutants TYMC-noCP, TYMC-Δ5707–6062, and TYMC-5985^{fs} showed that viral RNA is able to spread from cell to cell in the absence of coat protein, inducing the formation of local lesions. In contrast, mutants of TYMC with interrupted ORF-69 expression failed to induce lesion formation even though the accumulations of viral products in protoplasts were similar to wild type (Bozarth *et al.,* 1992; Figs. 3A and 3C, lane 9 cf. Fig. 4, lane 9). ORF-69 expression is thus critical for local movement and the coat protein dispensable. The inability of TYMC-noCP, TYMC-Δ5707–6062, and TYMC-5985^{fs} RNAs to move out of inoculated leaves (Figs. 4 and 5) clearly demonstrates the importance of the coat protein in permitting efficient long-distance viral transport and the establishment of a systemic infection. As we report here for TYMV, coat protein expression is dispensable for cell-to-cell spread but plays an important role in long-distance transport for a number of other plant members of the alphavirus-like group, including TMV (Dawson *et al.,* 1988; Saito *et al.,* 1990), the bromoviruses (de Jong and Ahlquist, 1991), beet necrotic yellow vein virus (Quillet *et al.,* 1989), and potato virus X (Chapman *et al.,* 1992). For at least one member of this group, barley stripe mosaic virus, the coat protein is entirely dispensable for long distance movement of viral RNA (Petty *et al.,* 1990), while for tobacco rattle virus, the situation is between the two extremes: the absence of coat protein results in virulent infections supported by cell-to-cell, rather than phloem-mediated, viral spread (Harrison and Robinson, 1978).

Our experiments with TYMC-C107/G96 have demonstrated that a short addition to the C-terminus of the coat protein does not prevent virion assembly, although the resultant capsid appears to be less stable than normal and the genomic RNA is poorly protected. According to our current understanding of TYMV encapsidation, protein–protein interactions involving the central hydrophobic domain of the CP are primarily responsible for capsid assembly (Re and Kaper, 1975), while the N- and C-termini are located on the surface of the virions (Quesniaux *et al.,* 1983). Crosslinking studies have detected extensive RNA–protein interactions except in the C-terminal and central hydrophobic domains of the coat protein (Ehresmann *et al.,* 1980). Our finding that a short C-terminal extension to the CP does not prevent encapsidation is consistent with this model, and it may indeed be possible to produce bioactive peptides in plants as C-terminal fusions of the TYMC coat protein, as suggested for TMV (Takamatsu *et al.,* 1990).

Information density and evolution of the TYMV genome

The region of the TYMV genome that we have studied in this paper appears to be constrained in its evolution by the presence of overlapping ORFs, even though we have found the ORF-206 readthrough domain to be nonessential. The readthrough domain is the least conserved of all ORFs between Australian (Keese *et al.,* 1989) and European isolates of TYMV (Dreher and Bransom, 1992), indicating some sequence drift, yet the length of the ORF is conserved (several substitutions introducing a stop codon in the readthrough ORF but silent in CP-ORF are possible). Our studies do not indicate why the readthrough ORF exists or has been evolutionarily retained.

This part of the TYMV genome is unusually information-rich, since it additionally includes the subgenomic start site and presumably contains sequences that direct subgenomic transcription. The initiation site for subgenomic RNA even overlaps with the 3′-most codon of ORF-206 (Fig. 1), and indeed an overlap of this type exists

for all characterized tymoviruses (Ding *et al.,* 1990). An identical arrangement exists in Sindbis virus RNA, where the 26S subgenomic RNA initiation site overlaps the 3′-most codon of the polymerase nsP4 (Strauss *et al.,* 1984), which is analogous to TYMC p70 (although there is no readthrough domain beyond the nsP4 ORF). The significance of these overlaps is not known, but a common heredity in the evolution of subgenomic RNA synthesis is suggested among these alphavirus-like group members.

The remarkable extent of overlapping functions present in the TYMV genome may have arisen as a result of the encapsidation constraints imposed on the genome of a small icosahedral virus. The existence of ORF overlaps also means that the protein domains encoded by TYMV RNA are unlikely to have recently arisen by horizontal transfer involving recombination with a donor RNA (another viral genome or a host mRNA). Such recombinational shuffling is considered to have been a major factor in positive strand RNA virus evolution (Koonin and Dolja, 1993).

ACKNOWLEDGMENTS

We thank the Electron Microscope Facility, Oregon State University, for help with the electron microscope studies. This work was funded in part by NSF Grant DMB-9019174 and USDA Grant 91-37303-5855. This is Technical Report No. 9871 from the Oregon Agricultural Experiment Station.

REFERENCES

BEIER, H., BARCISZEWSKA, M., KRUPP, G., MITNACHT, R., and GROSS, H. J. (1984). UAG readthrough during TMV RNA translation: Isolation and sequence of two $tRNAs^{Tyr}$ with suppressor activity from tobacco plants. *EMBO J.* **3,** 351–356.

BOZARTH, C. S., WEILAND, J. J., and DREHER, T. W. (1992). Expression of ORF-69 of turnip yellow mosaic virus is necessary for viral spread in plants. *Virology* **187,** 124–130.

BRANSOM, K. L., WEILAND, J. J., and DREHER, T. W. (1991). Proteolytic maturation of the 206-kDa nonstructural protein encoded by turnip yellow mosaic virus RNA. *Virology* **184,** 351–358.

CHAPMAN, S., HILLS, G., WATTS, M., and BAULCOMBE, D. (1992). Mutational analysis of the coat protein gene of potato virus X: Effects on virion morphology and viral pathogenicity. *Virology* **191,** 223–230.

CHEN, E., and SEEBURG, P. H. (1985). Supercoil sequencing: A fast and simple method for sequencing plasmid. DNA *DNA* **4,** 165–170.

CHURCH, G. M., and GILBERT, W. (1984). Genomic sequencing. *Proc. Natl. Acad. Sci. USA* **81,** 1991–1995.

DAWSON, W. O., BUBRICK, P., and GRANTHAM, G. L. (1988). Modifications of the tobacco mosaic virus coat protein gene affecting replication, movement and symptomatology. *Phytopathology* **78,** 783–789.

DE JONG, W., and AHLQUIST, P. (1991). Bromovirus host specificity and systemic infection. *Semin. Virol.* **2,** 97–105.

DING, S. W., HOWE, J., KEESE, P., MACKENZIE, A., MEEK, D., OSORIO-KEESE, M., SKOTNICKI, M., SRIFAH, P., TORRONEN, M., and GIBBS, A. (1990). The tymobox, a sequence shared by most tymoviruses: Its use in molecular studies of tymoviruses. *Nucleic Acids Res.* **18,** 1181–1187.

DREHER, T. W., and BRANSOM, K. L. (1992). Genomic RNA sequence of turnip yellow mosaic virus isolate TYMC, a cDNA-based clone with verified infectivity. *Plant Mol. Biol.* **18,** 403–406.

EHRESMANN, B., BRIAND, J. P., REINBOLT, J., and WITZ, J. (1980). Identification of binding sites of turnip yellow mosaic virus protein and RNA by crosslinks induced *in situ. Eur. J. Biochem.* **108,** 123–129.

GOLDBACH, R., LE GALL, O., and WELLINK, J. (1991). Alpha-like viruses in plants. *Semin. Virol.* **2,** 19–25.

GUILLEY, H., and BRIAND, J. P. (1978). Nucleotide sequence of turnip yellow mosaic virus coat protein mRNA. *Cell* **15,** 113–122.

HARRISON, B. D., and ROBINSON, D. J. (1978). The tobraviruses. *Adv. Virus Res.* **23,** 25–77.

ISHIKAWA, M., MESHI, T., OHNO, T., and OKADA, Y. (1991). Specific cessation of minus-strand RNA accumulation at an early stage of tobacco mosaic virus infection. *J. Virol.* **65,** 861–868.

KEESE, P., MACKENZIE, A., and GIBBS, A. (1989). Nucleotide sequence of the genome of an Australian isolate of turnip yellow mosaic virus. *Virology* **172,** 536–546.

KOONIN, E. V., and DOLJA, V. V. (1993). Evolution and taxonomy of positive-strand RNA viruses: Implication of comparative analysis of amino acid sequences. *Crit. Rev. Biochem. Mol. Biol.* **28,** 375–430.

LANE, L. C. (1986). Propagation and purification of RNA plant viruses. *Methods Enzymol.* **118,** 687–696.

MARSH, L. E., HUNTLEY, C. C., POGUE, G. P., CONNELL, J. P., and HALL, T. C. (1991). Regulation of (+):(−) strand asymmetry in replication of brome mosaic virus RNA. *Virology* **182,** 76–83.

MATTHEWS, R. E. F. (1991). "Plant Virology," 3rd ed. Academic Press, New York.

MORCH, M.-D., DRUGEON, G., ZAGORSKI, W., and HAENNI, A.-L. (1986). The synthesis of high molecular weight proteins in the wheat germ translation system. *Methods Enzymol.* **118,** 154–164.

MORCH, M.-D., BOYER, J.-C., and HAENNI, A.-L. (1988). Overlapping open reading frames revealed by complete nucleotide sequencing of turnip yellow mosaic virus genomic RNA. *Nucleic Acids Res.* **16,** 6157–6173.

PETTY, I. T., and JACKSON, A. O. (1990). Mutational analysis of barley stripe mosaic virus RNA β. *Virology* **179,** 712–718.

QUESNIAUX, V., JAEGLE, M., and VAN REGENMORTEL, M. H. V. (1983). Immunochemical studies of turnip yellow mosaic virus. Localization of two viral epitopes in residues 57–64 and 183–189 of the coat protein. *Biochem. Biophys. Acta* **743,** 226–231.

QUILLET, L., GUILLEY, H., JONARD, G., and RICHARDS, K. (1989). *In vitro* synthesis of biologically active beet necrotic yellow vein virus RNA. *Virology* **172,** 293–301.

RE, G. G., and KAPER, J. M. (1975). Chemical accessibility of tyrosyl and lysyl residues in turnip yellow mosaic virus capsids. *Biochemistry* **14,** 4492–4497.

SAITO, T., YAMANAKA, K., and OKADA, Y. (1990). Long-distance movement and viral assembly of tobacco mosaic virus mutants. *Virology* **176,** 329–336.

SARKAR, G., and SOMMER, S. S. (1990). The "megaprimer" method of site-directed mutagenesis. *BioTechniques* **8,** 404–407.

SKUZESKI, J. M., NICHOLS, L. M., and GESTELAND, R. F. (1990). Analysis of leaky viral translation termination codons *in vivo* by transient expression of improved β-glucuronidase vectors. *Plant Mol. Biol.* **15,** 65–79.

SKUZESKI, J. M., NICHOLS, L. M., GESTELAND, R. F., and ATKINS, J. F. (1991). The signal for a leaky UAG stop codon in several plant viruses includes the two downstream codons. *J. Mol. Biol.* **218,** 365–373.

SRIFAH, P., KEESE, P., WELLER, G., and GIBBS, A. (1992). Comparisons of the genomic sequences of erysimum latent virus and other tymoviruses: a search for the molecular basis of their host specificities. *J. Gen. Virol.* **73,** 1437–1447.

STRAUSS, E. G., RICE, C. M., and STRAUSS, J. H. (1984). Complete nucleotide sequence of the genomic RNA of Sindbis virus. *Virology* **133,** 92–110.

SUZUKI, M., KUWATA, S., KATAOKA, J., MASUTA, C., NITTA, N., and TAKANAMI, Y. (1991). Functional analysis of deletion mutants of cucumber mosaic virus RNA 3 using an *in vitro* transcription system. *Virology* **183,** 106–113.

TAKAMATSU, N., WATANABE, Y., YANAGI, H., MESHI, T., SHIBA, T., and OKADA, Y. (1990). Production of enkephalin (fusion protein) in tobacco protoplasts using a tobacco mosaic virus RNA vector. *FEBS Lett.* **269,** 73–76.

TSAI, C.-H., and DREHER, T. W. (1992). Second-site suppressor mutations assist in studying the function of the 3′ noncoding region of turnip yellow mosaic virus RNA. *J. Virol.* **66,** 5190–5199.

VAN DER KUYL, A. C., NEELEMAN, L., and BOL, J. F. (1991). Role of alfalfa mosaic virus coat protein in regulation of the balance between viral plus and minus strand RNA synthesis. **185,** 496–499.

VRATI, S., MANN, D. A., and REED, K. C. (1987). Alkaline Northern blots: transfer of RNA from agarose gels to Zetaprobe membrane in dilute NaOH. *Bio-Rad Molec. Biol. Reports* **1,** 1–4.

WEILAND, J. J., and DREHER, T. W. (1989). Infectious TYMV RNA from cloned cDNA: Effects *in vitro* and *in vivo* of point substitutions in the initiation codons of two extensively overlapping ORFs. *Nucleic Acids Res.* **17,** 4675–4687.

WEILAND, J. J., and DREHER, T. W. (1993). *Cis*-preferential replication of the turnip yellow mosaic virus RNA genome. *Proc. Natl. Acad. Sci. USA* **90,** 6095–6099.